Simplifying APACHE II Scoring using PDA’s

Authors:
Ian Galloway, B.Sc., RRT
Mary-Gordon MacKenzie, PhD
Andrew McIvor, MD, FRCP(C), FCCP
Paul Hernandez, MDCM, FRCP(C)

Institution at which the work was performed:
Capital District Health Authority
Halifax, Nova Scotia,
Canada, B3H3A7

Introduction:
The acute physiology and chronic health evaluation (APACHE II) is a scoring system that provides a means for describing and predicting acute illness severity over a broad range of intensive care unit (ICU) patients. Clinical investigators commonly use the APACHE II score in their research protocols to identify differences between treatment groups. Severity scoring systems, such as the APACHE II, are also routinely used in ICU’s as part of the audit and management process and for monitoring the acuity of patients within the ICU for quality assurance purposes.

Despite the usefulness of the APACHE II scoring system, previous research has documented barriers to its accurate recording. Polderman, Jorna and Gribes demonstrate that the interobserver variability in APACHE II scoring decreased when strict guidelines and regular training were implemented. These researchers also noted that even after the implementation of guidelines and a training program, calculation error remained a persistent source of variability in the scoring process.

The computation of APACHE II scores requires large amounts of data to be reviewed and analyzed. There are several complex pre-calculationsthat are required even before weighting and summing of the predictor variables. When done manually, the complexity of these multiple tasks introduce the possibility of frequent error. Gooder, Farr, & Young noted within their ICU, as a result of staff time constraints and the time required completing an APACHE II score, a large portion of ICU patients scores were not completed for 2 to 8 weeks. Using their hospital’s computerized patient health records and a custom-built APACHE II computerized scoring system, Gooder et al were able to automate the scoring of APACHE II scores. This created a more efficient system with timely reporting of the scores and a reduction in calculation errors.

To take full advantage of an automated APACHE II scoring system, the necessary infrastructure must be in place. Currently, in Canada, much of the clinical and administrative information in the health care system is maintained in files of paper records. The cost of a computerized patient health record system along with an automated APACHE II scoring system is a limiting factor for many ICU’s. Less expensive alternatives for computerized APACHE II scoring have been designed for desktop computers and the Internet. Although helpful, the use of these applications is dependent on the proximity of a desktop computer to the patient data.
As a result of the described complexity in scoring an APACHE II and the observed benefit of a computerized scoring system, we decided to look at personal digital assistants (PDA’s), which would give the user the ability to bring the computer to the patient’s bedside, or anywhere the patient data may be kept. We reasoned that collecting data using a custom built APACHE II software application (A2S) for a PDA in conjunction with the implementation of guidelines and a training program could decrease the variability associated with APACHE II scoring by decreasing the persistent calculation error inherent in the scoring system. The purpose of this study is to describe the results of the validation process used for the A2S custom software.

Methods:

A retrospective analysis of one hundred (n = 100) randomly selected APACHE II scores were recalculated. Individual APACHE II scores were identified from a previous research study conducted at the Capital District Health Authority in Halifax, Nova Scotia, Canada. An expert scorer, who had received training on scoring techniques using APACHE II guidelines, had compiled the original scores using a standard paper-scoring sheet. All the values for the APACHE II predictor variables from the previous research study had been recorded and saved on study specific forms.

Scores captured on these forms showed each of the predictor variables that had produced the highest score based on the patient’s highest or lowest derangement from the expected normal range for the first 24 hours of admission to the ICU. Also captured on these forms were the worst arterial blood gas results, systolic and diastolic pressure readings, age, and co-morbid condition. Calculations for the mean arterial blood pressure, alveolar to arterial (A-a) gradient, Acute Physiological Score (APS), age points, and the final APACHE II score for had also been recorded. Values for the chronic health points were also recorded; however no detailed information was noted for the assignment of the chronic health points.

The previously recorded patient data was inputted into the A2S to obtain a new APACHE II score. Manual recalculation using a standard paper-scoring sheet acted as the control score. The standard paper-scoring sheet has been empirically shown to be an accurate method for scoring APACHE II scores. The objective of our testing was to determine if there was an overall agreement in the APACHE II scores found between the three scoring methods. When the scores differed, a detailed review comparing the A2S, the standard paper-scoring sheet and the expert scorer was conducted to reveal which factors attributed to the discrepancy.

Materials:

In 1996 Palm released the Pilot 1000 series (Palm Inc; Milpitas, CA) of the PDA commonly referred to as a Palm. The introduction of this handheld computing device was an evolutionary step forward from calculators and date books to actual pocket computers. Since 1996, Palm’s processing power, memory and operating system (OS) have improved significantly. Additional characteristics of PDA’s, such as, ease of use, simplicity to synchronize with desk top and network computer systems, portability, and affordability have contributed to their increase in popularity.
The Palm’s OS has set the standard for PDA operating systems. Palm OS can be found in a variety of brand name PDA’s (e.g. Sony, Symbol, Handspring, Samsung, Kyocera and IBM). As a result of the wide availability of Palm OS PDA’s, and the large number of physician now using them, we decide to build the A2S for the Palm OS.

**Hardware:**

All testing was done using the Handspring Visor Platinum (HandSpring Inc; Mountain View, CA) PDA. The Handspring Visor Platinum utilizes a 33 MHz Dragonball VZ CPU (Motorola Inc; Schaumburg, Ill) processor with 8 megabytes of random access memory (RAM), Palm 3.5 OS and has a 160x160-pixel monochrome screen.

**Palm OS software program (Apache.prc):**

There are a great number of development environments for creating Palm OS applications. Metrowerks CodeWarrior 20, for Palm computing platform, was selected as the integrated development environment (IDE) to develop the A2S. CodeWarrior, which uses the C and C++ programming language, contains all the necessary tools required to develop Palm OS applications.

The CodeWarrior IDE is a multiple document interface that allows the user to work on more than one component of the application at a time (e.g. source code and/or the graphical user interface). CodeWarrior organizes the application into a single project file, which makes reference to all the various components that make up the application (e.g. source code, resource files, and libraries). Once the project is completed CodeWarrior compiles the source code (textual programming statements) creating an object file linking it to the libraries that define the inner workings of the handheld device. In addition the resource files (textual description of the graphic user interface) are compiled and linked to the object files and libraries. The final product of this process is a Palm resource database file known as a PRC file (PRC is also the file extension), which can be run in any Palm OS device as an executable software application.

The primary goal in the development of the A2S was to eliminate the burden of calculation embedded within the APACHE II scoring procedure. In order to simplify the interface, the user was required to only input the highest or lowest value for each of the 12 predefined predictor variables, the patient’s age and the assigned value for the Chronic Health Points (CHP). The software would then calculate the core temperature, mean arterial pressure, A-a gradient, weight the predictor variables (on a scale from 0 to 4) and calculate the acute physiology score (APS).

Once all the physiological measurements, CHP and age points have been entered. The user then taps on the "Calculate Score" button and the application displays a new form with a calculated APACHE II score. This new form allows the user to select a detailed break down of the score by tapping on the "Detailed Results" button.

The detailed results are displayed in a new form showing the values inputted and the weight assigned to each predictor variable, age point and CHP. If a change is required for any inputted value, the user simply taps on that value and the associated form for
that variable reappears. This allows the user to make any necessary changes and permit a recalculation of the score, if required.

Results:

Using the standard paper-scoring sheet and the A2S, two new APACHE II scores were recalculated using the APACHE II predictor variables data taken from the previous studies forms. All three scores were compared. When the scores differed, the investigator conducted a detailed review of each score to determine where the error(s) may have occurred.

Eighty-one (81%) of the scores where identical when comparing the expert scores, A2S and the standard paper-scoring sheet while nineteen (19%) scores differed. Detailed analysis of these 19 scores showed that the standard paper-scoring sheet and the A2S were in agreement 100% of the time. The scoring disagreement was found to be the result of miscalculations made by the expert scorer.

Mathematical error and inaccurate weighting were the main causes of the errors (Table 1). Weighting of the variables was the greatest source of error accounting for 78% (n = 15), errors in summing of the variables was responsible for 11% (n = 2) and the remaining 11% (n = 2) were due to pre-calculation errors.

APACHE II scores ranged from 0 to 37 (Table 2). Comparing the 19 scores that differed, 9 (47%) of the scores varied by 1 point, 8 (42%) varied by 2 points, 1 (5.5%) varied by 3 points and 1 (5.5%) varied by 5 points. The A2S and the standard paper-scoring sheet values were higher than the expert scorer’s APACHE II results for 14 of the 19 scores (74%). The expert scorer’s APACHE II results were higher in 5 of these 19 scores (26%).

Discussion:

The goal for this study was to validate the A2S custom software by comparing it to a manually scored APACHE II. As stated previously, the complexity of the APACHE II scoring system increases the likelihood of human error. Presently the gold standard for APACHE II scoring is the manual paper-scoring sheet. However the 19% error rate we found when using a manual paper-scoring sheet, demonstrates the APACHE II’s complexity and propensity to error even when scored by an expert. These findings are consistent with those reported by Polderman et al., and demonstrate that calculation error is inherent in the APACHE II scoring system. Utilizing the A2S we were able to produce results that were more accurate than the manual scoring of the APACHE II done by an expert.

The A2S is designed to eliminate the burden of calculation embedded within the APACHE II scoring process. Small errors made when calculating the pre-APACHE II calculations, weighting the variables or summing the score can lead to an incorrect APACHE II score. Detailed analysis of the nineteen varied scores showed that mathematical error and inaccurate weighting were the main causes for errors made by the expert scorer. Weighting of the variables was the greatest source of error at 78% mathematical error accounted for 22%. It is not clear in this study why one type of error was more prevalent.
A single APACHE II score is generally not of clinical significance and should not be used as a triaging tool. APACHE II scores are best used for describing and predicting illness severity over a broad range of ICU patients. When averaged, the APACHE II becomes a useful tool to identify differences among groups of ICU patients. However, if persistent small calculation errors are not captured, they may lead to incorrect characterization of an ICU patient population.

When an error was made, the expert scorer’s value was usually lower than that obtained by the A2S. The reason for this is unclear and has been noted in earlier research comparing experts manually scoring APACHE II scores versus a computer calculated scores. Further investigation would need to be conducted to determine if this is indeed true or just a chance occurrence.

There are some potential limitations of this study design. The most significant limitation is that only one person calculated the original 100 scores. It is not clear if other individuals would produce more or less accurate APACHE II scores. A larger study to assess the accuracy of different users would be needed to determine if the error rate change with more scorers. Studying a larger group would also allow us to determine if there is a user preference for scoring an APACHE II. This study did not assess the potential time saved by using the software, as manual APACHE II scoring is a laborious process. A comparison of the time required calculating an APACHE II score by using the A2S versus the standard paper-scoring sheet may prove to be an added value to the usefulness of the A2S.

In conclusion, the A2S may help decrease the incidence of computational errors and improve consistency among scorers as compared with a manually scored APACHE II using a standard paper-scoring sheet. This may prove helpful in an attempt to improve scoring accuracy in research studies and clinical audits. Future research is needed to assess larger scale use of the A2S in the ICU in terms of scorer preference, time saving potential and accuracy when applied to multiple users.

References:

5. Garrard CS. Human-computer interactions: can computers improve the way doctors work? Schweiz Med Wochenschr Suppl 2000; 130(42): 1557-1563
collection of the APACHE II scores in teaching and community hospitals Crit
Care Med 1999; 27(9): 1999-2004
Accuracy and efficiency of an automated system for calculating APACHE II
scores in an intensive care unit. Available at:
9. Holt AW, Bury LK, Bersten A D, et al. Prospective evaluation of residents and
APACHE II scoring. Intensive Care Med 2001; 27(9): 1550-1552
11. Polderman KH, Girbes AR, Thijs LG, et al. Accuracy and reliability of
APACHE II scoring in two intensive care units. Anaesthesia 2001; 56(1): 47 -50
12. Romanow RJ. Commission on the Future of Health Care in Canada (Final
Report). Subset - Information, Evidence and Ideas. Available at:
http://www.healthcarecommission.ca/pdf/HCC_Chapter_3.pdf Accessed:
December 31, 2002
13. Edwards FH, Davies RS. The handheld computer in critical care medicine. Am
Surg 1986; 52(8): 452-455
15. Palm Inc. Product information 2001. Palm and ePocrates Team To Deliver
Custom Handheld Solution For Healthcare Professionals Available at:
16. Perahia A. Suny Downstate Medical Center 2002. The Personal Digital
Assistant: A Doctors best friend. Available at: http://ect.downstate.edu/pda
Accessed: March 22, 2003
17. Rowin Group. Participating in e-marketing - The ePharm 5th from
ePhamaceuticals and iPphysiciansNet. More Physicians using PDA’s. Available
22, 2003
computing devices increases 15% in 1999 to 26% in 2001. Available at:
Accessed: March 22, 2003
19. Handspring Inc. Product information web site. Available at:
March 11, 2002
Available at: http://www.metrowerks.com/products/palm Accessed: Aug 13,
2002
2000; 41-97